African Journal of Biotechnology Vol.11(23), pp. 6226-6233, 20 March, 2012
Available online at http://www.academicjournals.org/AJB
DOI: 10.5897/AJB11.4117
ISSN 1684–5315 © 2012 Academic Journals

Full Length Research Paper

Analysis of the copy number of exogenous genes in transgenic cotton using real-time quantitative PCR and the $2^{-\triangle\triangle CT}$ method

Xiaojie Yang[1,2], Fuguang Li[1]*, Chuanliang Liu[1], Xueyan Zhang[1], Kun Liu[1], Weiping Fang[2], Zhixia Wu[1], Deyi Xie[2], Chaojun Zhang[1], Qianhua Wang[1] and Fu′ an Zhao[2]

[1]State Key Laboratory of Cotton Biology, Cotton Research Institute, Chinese Academy of Agriculture Sciences, Anyang, 455000, China.
[2]Economic Crop Research Institute, Henan Academy of Agriculture Sciences, Zhengzhou 450002, China.

Accepted 23 February, 2012

Presence of multiple copies of a transgene has been found to associate with gene silencing that may manifest early or only over a period of time. As such, transgenic copy number should be analyzed as soon as possible. The commonly used method to analyze gene copy number, Southern blotting, has been found to be unreliable for determination of gene copy number when rearrangement or tandem repeats integration occurred. In recent years, a powerful real-time fluorescence quantitative real-time PCR (qRT-PCR) method has been used to analyze gene copy number in genetically modified plants, but these methods have also their own application conditions. Thus, based on real-time quantitative PCR technology, the modified $2^{-\triangle\triangle CT}$ method, which has been used usually in analyzing gene expression, was used for the first time, to analyze transgenic copy numbers in transgenic cotton. In this paper, a single-copy homozygous transgenic plant whose identity was confirmed by Southern blotting using restriction enzymes was used as a calibrator. Genomic DNA that was extracted from the calibrator and six transgenic T_0 cotton plants transformed by pollen tube pathway or gene gun bombardment was amplified by quantitative PCR, and the gene copy numbers were estimated by the modified $2^{-\triangle\triangle CT}$ method to be 6, 3, 2, 2, 3 and 2 copies. In addition, except 6 copies all other identifications (3, 2, 2, 3 and 2) were equally indicated by the corresponding Southern blotting. This study supports the modified $2^{-\triangle\triangle CT}$ method using quantitative real-time PCR technology to provide a fast, high-throughput method to analyze the copy number of foreign genes in a large transgenic T_0 population.

Key words: Cotton, transgene copy number, modified $2^{-\triangle\triangle CT}$ method, Southern blotting.

INTRODUCTION

Expressing a transgene can be affected by a number of factors such as site of insertion, copy number, and the direction of adjacent transgene copies for example head to head or head to tail. Higher copies of a transgene can cause co-suppression leading to transgene silencing. All of the commonly used genes transfer protocols, namely *Agrobacterium* mediation, particle gun bombardment, and pollen tube pathway injection can produce multiple gene copies. Generally, higher copies of transgene are found more often in plants obtained by biolistic bombartment and pollen tube injection and less in plants obtained by *Agrobacterium*-mediated protocol. Even so, the trans-

*Corresponding author. E-mail: lifug@cricaas.com.cn; aylifug@hotmail.com. Tel: +86-372-2562204. Fax: +86-372-2562256.

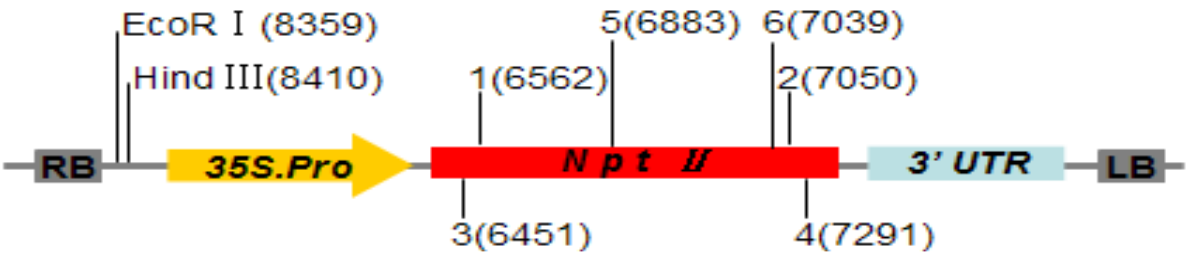

Figure 1. Schematic map of T-DNA region with expression vector pCAMBIA 2300 for transformation. The RB and LB represent the right border and left border, respectively. 1 to 2 and 3 to 4: represent the primers (sites) for molecular identification of positive transgenic cotton plants and southern blot probe synthesis, respectively; 5 and 6 represent the primers for real-time PCR analysis.

formants derived from *Agrobacterium* mediation may still contain three or more copies of transgene. Gene silencing induced by multiple gene copies may manifest early after the transformation. It may also be delayed fora period of time or even until next or subsequent generations. As such, indefinite time will be needed to ascertain whether the expression of a transgene is stable over time. As there is a correlation between multiple gene copies with transgene silencing, early detection of transgene copies can be useful in sorting out the transgenic plants with higher tendency of transgene silencing.

A number of methods are available for analyzing transgene copy number; these include Southern blotting, comparative genomic hybridization (Kallioniemi et al., 1996), multiplex probe amplification and hybridization (Armour et al., 2000), and chip hybridization (Lucito et al. 2000). There are a number of shortcomings, for example laborious protocol, requirement of large amounts of DNA, and inaccurate estimation of the foreign gene copy number when tandem integration or rearrangement of the foreign gene occurs associated with these methods (Mason et al., 2002). In recent years, a powerful real-time fluorescence quantitative real-time PCR (qRT-PCR) method has been used to analyze gene copy number in transgenic corn, rapeseed, rice, and cotton plants (Song et al., 2002; Weng et al., 2004; Yang et al., 2005a; Yang et al., 2005b). This method does require certain application conditions, such as the selection and copy number identification of the endogenous gene, accurate determination of DNA concentration. In this paper, the $2^{-\triangle\triangle CT}$ method, based on qRT-PCR and usually used in gene expression analysis (Pfaffl, 2001) was applied, for the first time, to analyze the T-DNA copy number of transgenic cotton, and the results were compared with those obtained by Southern blotting.

MATERIALS AND METHODS

Plasmid

Expression vector pCAMBIA2300 (Figure 1) kindly provided by Dr. Yongjun Lin was used as a positive control in Southern blotting analysis in this study.

Cotton plants (lines)

Upland cotton variety CCRI 24 was used in this study. A stable transgenic cotton line transformed by *Agrobacterium* tumefaciens strain LBA4404 containing the vector pCAMBIA 2300 was used as the calibrator in transgene copy number analysis. Four T_0 transgenic cotton lines transformed by pollen tube pathway and two T_0 transgenic cotton lines by gene gun bombardment were studied.

Chemical reagents

Kanamycin was obtained from Sangon Biotech (China). DNA markers and Taq polymerase from TianGen Biotech (China), and electrophoresis-grade agarose from Amresco (US) were used in screening positive transgenic cotton lines. The quantitative PCR kit from TaKaRa (China) were used in quantitative analysis. PCR product extraction kit from TianGen Biotech (China), restriction enzymes from Promega (US), the DIG Probe DNA Labeling and Detection Starter I kit from Roche (Switzerland), and the Hybond-N^+ nylon membrane from Pharmacia (Italy), were used in Southern blotting analysis.

Detection of kanamycin resistance

Kanamycin resistance was carried out by smearing kanamycin (1500 mg/l) on the leaf surfaces of three to five leaf stage transgenic cotton plants. After five to seven days, the leaves of non-transform plants turned brown whereas the leaves of transformed plants displayed no change.

Extraction of cotton genomic DNA and PCR identification for selectable *Npt*II gene

Leaves from kanamycin-resistant T_0 generation plants were wrapped in a plastic membrane and frozen in liquid nitrogen. Genomic DNA was extracted by a modified cetyl trimethy-lammonium bromide (CTAB) method (Paterson, 1993). Primers (forward primer: 5'-GCCATGTGGGCATTCAGT-3' and reverse primer: 5'-TACGCCATCAGCACGTTATCG-3') (Figure 1) were designed to amplify a coding region fragment of the *Npt*II(kanamycin resistance) gene. PCR reaction mixture (20 µl) consisted of 10 to 20 ng cotton genomic DNA, 0.2 mM dNTPs, 0.6

μM of each primer, 1 × PCR buffer and 0.15 U Taq. Thermal cycling was carried out at 94°C for 10 min, followed by 32 cycles of 94°C for 30 s, 58°C for 30 s, and 72°C for 30 s. After the final cycle, reaction was maintained at 72°C for 10 min before completion.

Restriction enzyme digestion of genomic DNA and Southern blotting

The genomic DNA *Eco*RI and *Hind*III restriction enzyme digestion system contained 30 μg of genomic DNA, 40 μl 10 × H buffer (10 × T buffer), 40 μl *Eco*RI (*Hind*III) (8 to 20 U/μl), and sterile ddH_2O to a final volume of 400 μl. The mixture was vortexed briefly to ensure complete mixing. Digestion was performed at 37°C. The PCR (forward primer: 5′-ACGCGACGCTTAGCCCTCGCCGC-3′ and reverse primer: 5′-GTTCTACCAACGGTGCGTCCA-3′) fragment of the *Npt*II gene was used as a probe for Southern blotting. Probe labeling and hybridization were performed according to the instructions for the DIG Probe DNA Labeling and Detection Starter kit I.

Establishment of standard curves for the endogenous reference gene *Sad*I and selectable marker *Npt*II gene

Primer 5.0 software was used to design the primer pairs for the selection marker gene NptII and the upland cotton *Sad*I gene (Yang et al., 2005b). The primers were: *Sad*I forward, 5′-CGGTCAGTGGAAAGGTAAGG -3′ and reverse, 5′-ACCAGCAAAAGGAATCAAGC-3′; *Npt*II forward, 5′-GATGTTTCGCTTGGTGGTCG -3′ and reverse, 5′-GGAAGGGACTGGCTGCTATTG -3′. The template DNA (100,ng/μl for *Sad*I and 200,ng/μl for *Npt*II) from the calibrator plant containing one copy of the exogenous gene was confirmed by Southern blotting and was diluted at a five serial concentration gradient. Each reaction was repeated thrice to establish standard curves for qRT-PCR. The 25 μl qRT-PCR reaction mixture contained the following components: 2 μl (100 ng) template DNA, 12.5 μl 2 × SYBR® Premix Taq™, 0.5 μl (10 μM) of the forward and reverse primers, 0.5 μl of the 50 × ROX Reference Dye II, and 9.0 μl of sterile ddH_2O. The qRT-PCR program included three steps, and the fluorescence signal was measured in the extended phase. When the PCR reaction was completed, the solubility curve was added. The PCR program consisted of 94°C for 10 min, 40 cycles of 94°C for 30 s, 56°C for 10 s, 72°C for 30 s, and 72°C for 10 min for the final extension. All reactions were performed with an ABI7500 quantitative real time PCR machine.

Analysis of copy number in the transgenic lines by modified $2^{-\Delta\Delta CT}$ method

Each qRT-PCR reaction contained DNA (50 to 100 ng) of transgenic cotton T_0 plants and a single-copy control. The endogenous reference and selection marker genes were amplified simultaneously, and each reaction was repeated three times. The C_T values obtained were used as the original data to estimate the copy number of the transgene. Copy number was calculated by the modified $2^{-\Delta\Delta CT}$ formula: copy number = 2 × ($E_N^{\triangle CTN\ (control\text{-}sample)}$) / ($E_S^{\triangle CTS(control\text{-}sample)}$), where N represents *Npt*II gene, S represents endogenous *Sad*I gene, and E is the amplification efficiency. The cycle threshold (C_T) value, also known as the threshold cycle, represents the required number of cycles for the fluorescence signal to increase above the baseline in qRT-PCR.

RESULTS

Kanamycin resistance test and PCR identification

Expression of *Npt*II gene in transgenic cotton plants from *Agrobacterium*-mediated and pollen tube pathway transformation was confirmed by the tolerance to kanamycin treatment (Figure 2B) by the leaves compared to the browning of the leaves of the non-transformed plants (Figure 2A). Presence of the *Npt*II gene was also confirmed by the presence of the specific 489 bp band of the *Npt*II sequence following PCR amplifications (Figure 2C).

Analysis of copy number of positive control transgenic cotton (used as the calibrator for analyzing copy number by $2^{-\Delta\Delta CT}$ method) by Southern Blotting

Southern blotting analysis of the calibrator plant showed a sharp band with the dual-single restriction enzymes (*Eco*RI and *Hind*III) digestion indicating the presence of a single copy of the *Npt*II gene (Figure 3).

Amplification and dissociation curves for *Sad*I and *Npt*II gene in the qRT-PCR amplification

Genomic DNA from the control transgenic plant confirmed by Southern blotting was diluted for concentration gradient (200, 40, 8.0, 1.6, and 0.32 ng/μl) to be used as templates for real-time PCR. Primers for the *Sad*I and *Npt*II gene were used for PCR amplification, and all amplifications were repeated thrice to produce qRT-PCR standard curves for the two genes. The amplification curve showed that the C_T value increased proportionally to the increase in each template concentrations, and showed a good reproducibility (Figures 4A and B), which indicated that the C_T values and concentrations could be used to generate standard curves. Both the dissociation curves also showed a single peak dissociation curve (Figures 5A and B), which indicated PCR amplifications were very specific, and had no primer dimers and nonspecific amplification.

qRT-PCR amplification efficiency for *Sad*I and *Npt*II gene

On the basis of $LogC_0$ (template initial concentration in qRT-PCR reaction) and linear C_T values, standard curve equations for *Sad*I and *Npt*II was obtained: $Y_{SadI} = -3.24x$

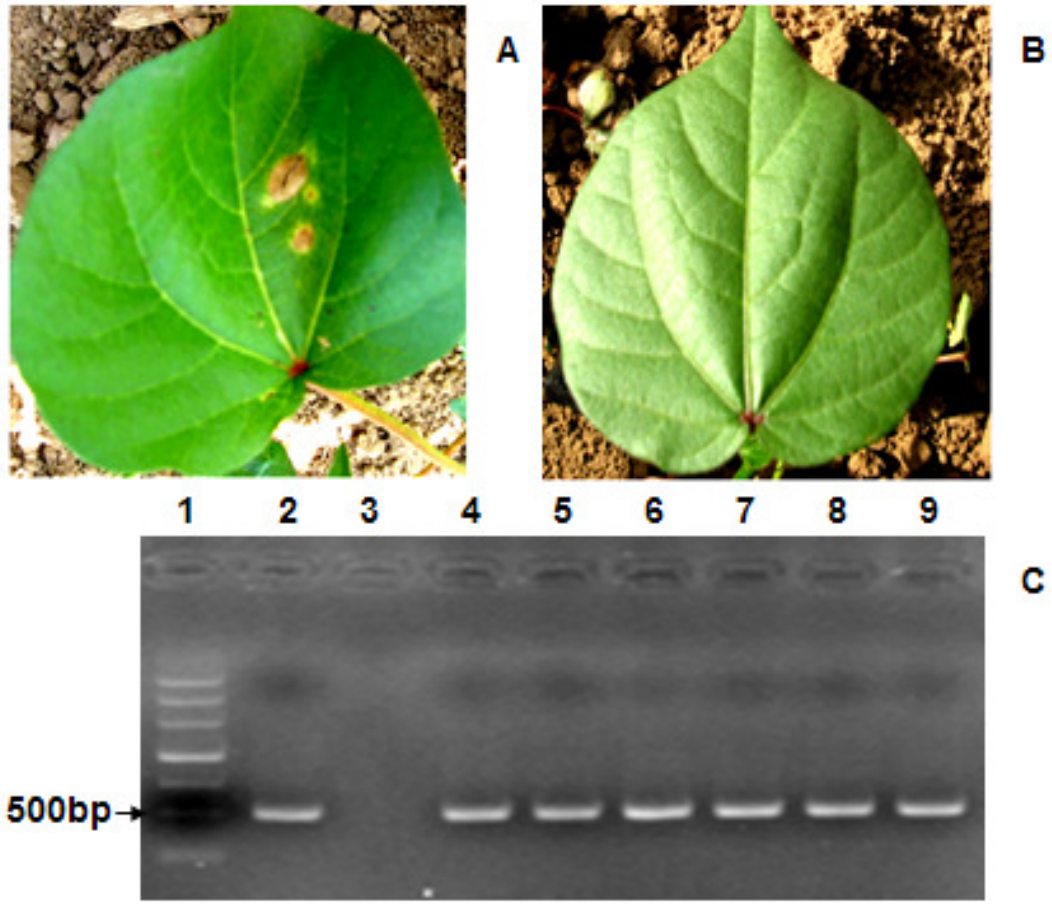

Figure 2. Kanamycin-resistance of transgenic cotton leaves and PCR identification of selectable marker *Npt*II gene. A and B: Kanamycin-sensitive and -resistant cotton plants, respectively; C: a conventional PCR of *Npt*II. 1, Marker iii; 2, positive plasmid control; 3, non-transgenic plant; 4 to 7 and 8 to 9, transgenic plants transformed by pollen tube pathway and gene gun bombardment mediated transgenic plants, respectively.

+ 20.37; $Y_{NPTII} = -3.43x + 19.84$. The equation coefficients of determination (R^2) were 0.997 and 0.998 for *Sad*I and *Npt*II, respectively, indicating good reproducibility of the linear relationship (Figures 6A and B). Amplification efficiencies of both E_{SadI} and E_{NptII} were calculated according to the formula: $E = 10^{-1/\text{linear equation slope}}$. The *Sad*I and *Npt*II gene amplification efficiencies (E_{SadI} and E_{NptII} of 1.95 and 2.03, respectively) indicated near 100% amplification of the both genes.

Estimation of transgene copy number

Homogenization of each template in qRT-PCR reactions was achieved by amplification of the *Sad*I gene. Genomic DNA from single-copy genetically modified (GM) cotton control, and six randomly selected T_0 positive lines were used as PCR reaction templates. *Sad*I and *Npt*II gene primers were used in all the PCR reactions, and each reaction was repeated thrice. The amplification curves for *Sad*I and *Npt*II showed good reproducibility. The *Sad*I gene was amplified in the negative control and also it is an endogenous gene, while *Npt*II gene was not amplified (Table 1). The standard deviation of the C_T value for the *Sad*I gene was 0.03 to 0.17, while it was 0.09 to 0.46 for the *Npt*II gene. For the both genes, the variation coefficient was < 1, indicating good reliability of the C_T values. Thus, the data could be used to analyze the copy number of exogenous genes in transgenic plants (Table 1).

The foreign gene copy numbers in transgenic plants was estimated by the $2^{-\Delta\Delta CT}$ method and the provided formula. In this study, a transgenic line was selected with a single copy of the transgene confirmed by Southern blotting as the positive control (calibrator). Using the formula, the copy number of the foreign gene was calculated in six randomly selected transgenic T_0 generation plants to be 6, 3, 2, 2, 3 and 2, respectively (Table 2).

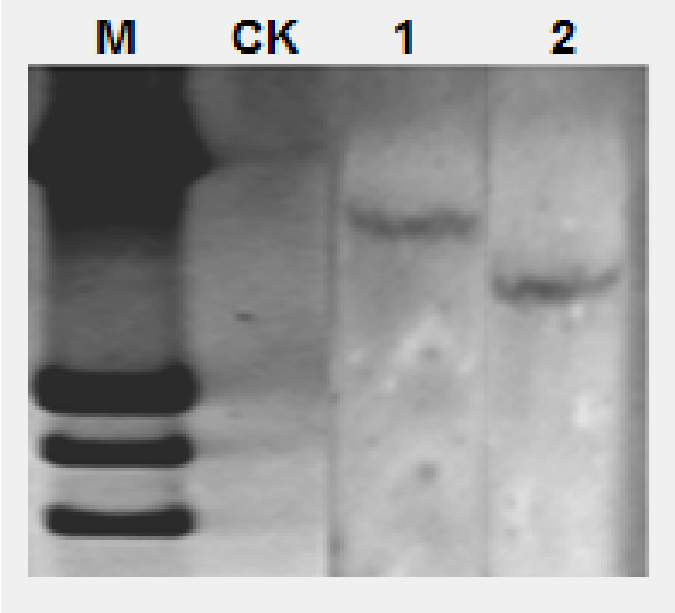

Figure 3. Southern blotting identification of DNA from homozygous transgenic cotton. Lane M, Marker iii; CK, non-transgenic plant; 1 and 2 represent the southern hybridization after dual single-enzyme *Eco*R I and *Hind* iii digestion of genomic DNA from homozygous transgenic cotton line.

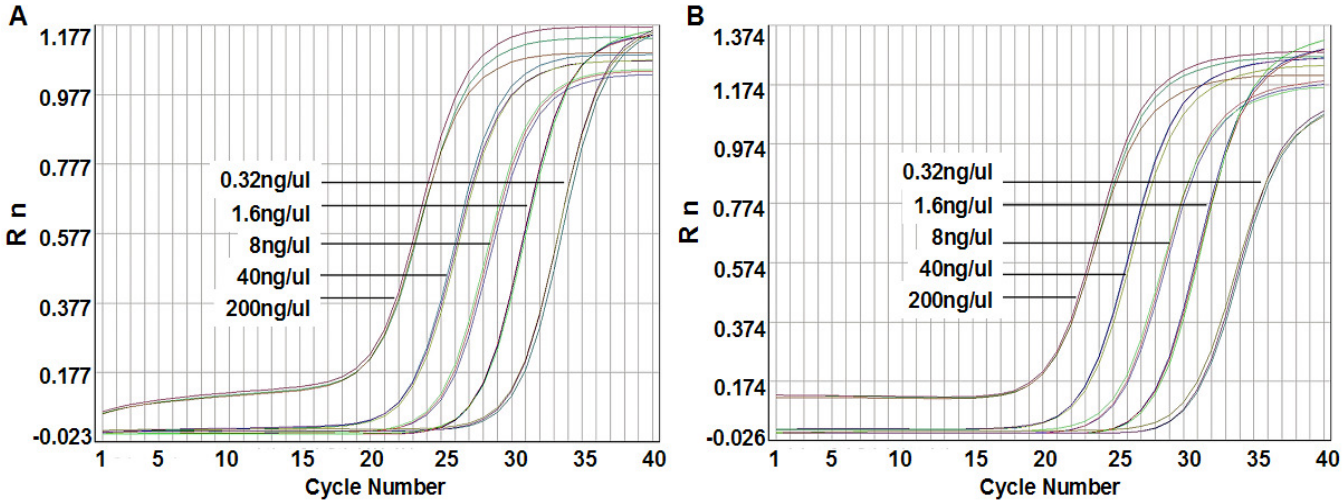

Figure 4. Amplification plots for the endogenous reference gene *Sad*I and selectable marker *Npt*II gene real-time PCR assays. The x-axis and y-axis represent real-time PCR cycle number and normalized reporter fluorescence (Rn), respectively. A, Amplification plot for *Sad*I gene; B, amplification curve for *Npt*II gene.

Southern blotting verification of the T_0 generation transgenic lines

To verify the estimation of exogenous gene copy number in the six GM cotton lines by the modified $2^{-\Delta\Delta CT}$ method, 80 µg plant genomic DNA from individual plants and subsequently were digested by *Eco*RI for Southern blotting analysis. The results show that five out of the six lines had similar results as that of the qRT-PCR results, that is, copy numbers 4, 3, 2, 2, 3 and 2, respectively. The results with quantitative PCR are either similar or show higher copy numbers with Southern blotting (Table

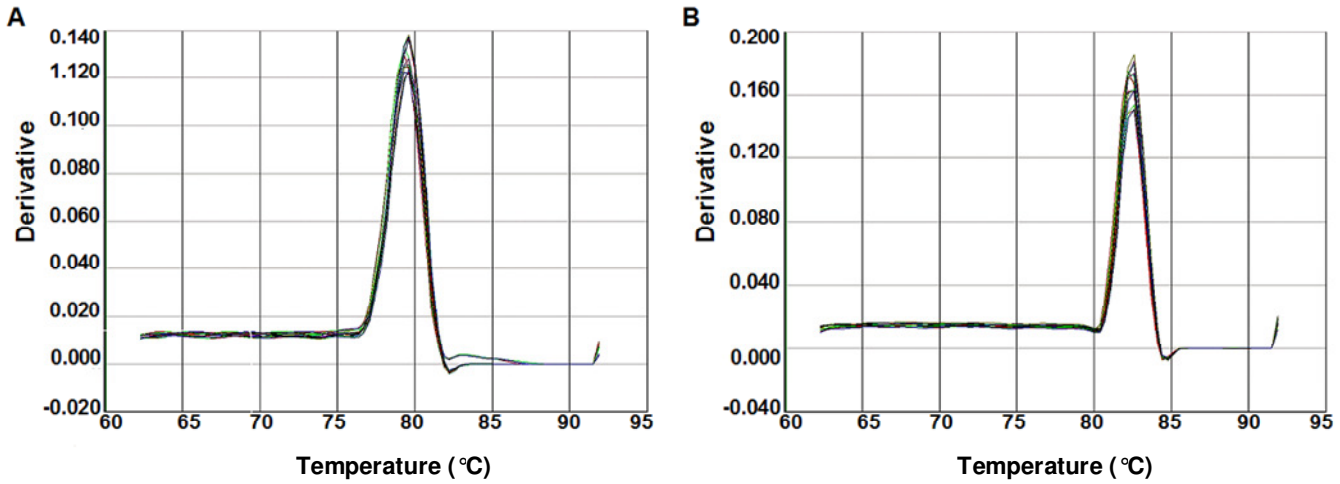

Figure 5. Dissociation curves for the endogenous reference gene *SadI* and selectable marker *Npt*II gene real-time PCR assays. The x-axis and y-axis represent temperature and the negative first derivative of fluorescence intensity (-dFI/dT), respectively. A, Dissociation curve for *Sad*I gene; B, dissociation curve for *Npt* II gene.

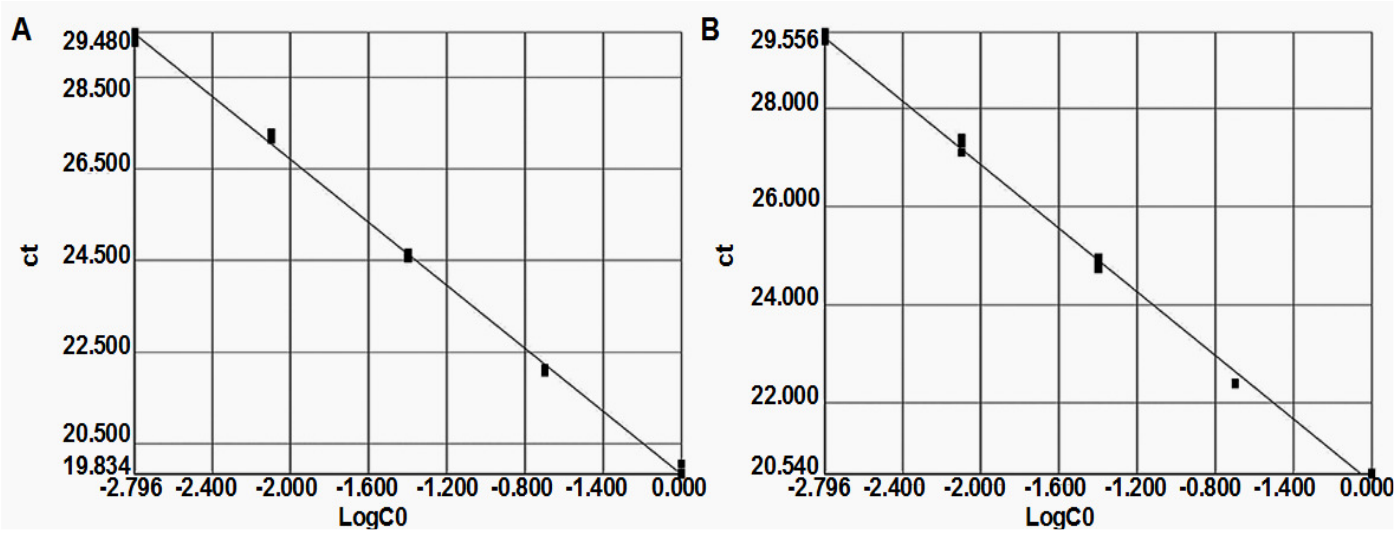

Figure 6. Standard curves for the endogenous reference gene *SadI* and selectable marker *Npt*II gene real-time PCR assays. The x-axis and y-axis represent the logarithm of each gradient initial concentration (logC0) and the threshold cycle(s) (Ct), respectively. A, Standard curve for *Sad*I gene; B, standard curve for *Npt*II gene.

2 and Figure 7).

DISCUSSION

In this paper, the modified $2^{-\Delta\Delta CT}$ method based on relative qRT-PCR used normally to analyze gene expression was used to analyze the transgene copy number in GM cotton. Due to the fact that the selected transgenic line used as the positive calibrator was homozygous and the possibility of homozygote of integrated foreign gene in the transgenic T_0 generation was very small, the ratio of $E_N^{\triangle CT\ N(control\ -\ sample)}$ to $E_S^{\triangle CT\ S(control\ -\ sample)}$ multiplied by 2 were the theoretical foreign gene copy numbers in the transgenic T_0 generation. The copy number of the transgene was calculated with the modified $2^{-\Delta\Delta CT}$ formula. The transgene copy numbers from the relative qRT-PCR method were found to be either similar or higher than those obtained by Southern blotting. Some early studies have shown that integrations of T-DNA tandem repeats, including forward and inverse repeats were a common phenomenon in GM plants (Kim et al.

Table 1. C_T values of samples based on the *Npt*II and *Sad*I gene standard curves.

Transgenic T_0 plant	*Npt*II		*Sad*I	
	C_T Value	CV (%)	C_T Value	CV (%)
1	20.72 ± 0.02	0.09	21.31 ± 0.04	0.21
2	20.46 ± 0.03	0.14	19.96 ± 0.10	0.52
3	22.32 ± 0.04	0.16	21.43 ± 0.16	0.73
4	21.52 ± 0.02	0.01	20.25 ± 0.02	0.08
5	20.69 ± 0.10	0.47	20.30 ± 0.06	0.30
6	21.80 ± 0.06	0.27	20.47 ± 0.15	0.72
Negative control	-	-	20.55 ± 0.04	0.10
Single-copy control	21.31 ± 0.02	0.38	20.30 ± 0.03	0.13

C_T value of each sample is expressed as the average ± standard deviation. CV, coefficient of variation for each group.

Table 2. Estimates of T-DNA copy number in transgenic plants.

Transgenic T_0 plant	$(E_N)^{\Delta C_T N}$	$(E_S)^{\Delta C_T S}$	$2\times(E_N)^{\Delta C_T N}/(E_S)^{\Delta C_T S}$	Copy number of transgene	
				q-PCR	Southern blot
1	1.5017	0.5081	5.91	6	4
2	1.7973	1.0556	3.41	3	3
3	0.4954	0.4680	2.12	2	2
4	0.8612	1.0348	1.66	2	2
5	1.5380	1.0018	3.07	3	3
6	0.7125	0.8904	1.6	2	2
Negative control	0.0000	0.3250	0.0000	0	-
Single-copy control	-	-	-	-	1

2003; Makarevitch et al. 2003; Zhang et al. 2008). Southern blotting may not exactly identify copy number in such plants. Therefore, presence of T-DNA tandem repeats may be the cause of the discrepancy in copy numbers between the qRT-PCR and the Southern blotting protocols.

Analyzing T-DNA copy number in transgenic plants using the modified $2^{-\Delta\Delta CT}$ method has many advantages. Firstly, this method could avoid estimating of the molecular weight of host plant genomic DNA, and estimating in itself will lead to a certain degree of error. At present, whole genomic sequences are available for rice, corn, sorghum, potatoes, rapeseed, and a few other crops. For other plants, the use of estimated genome DNA molecular weights for their genomes in absolute qRT-PCR analysis of exogenous gene copy number in plants may not give the accurate results. However, the modified $2^{-\Delta\Delta CT}$ method does not require the information for genome DNA molecular weight. The modified $2^{-\Delta\Delta CT}$ method does not also require information of the copy number of the endogenous reference gene. Earlier reports suggested that the information regarding copy number of the endogenous reference gene is a prerequisite for relative qRT-PCR analysis of transgene copy number (Weng et al., 2004). In modified $2^{-\Delta\Delta CT}$ analysis of transgenic cotton T-DNA copy numbers, endogenous reference genes only played a role realizing the genome DNA template homogenization in qRT-PCR, and so, identification of copy number of the endogenous reference gene is not necessary. Furthermore, with this method, exact measurement of DNA concentration can be avoided, reducing chances of error. Currently, other analysis method based on qRT-PCR of transgenic copy number is dependent on the DNA concentration, and even measurement of the volume of the reaction system (Mason et al. 2002), which can lead to errors during the analysis. The modified $2^{-\Delta\Delta CT}$ method effectively avoided errors caused by human measurement by using an internal reference gene in template homogenization, thus eliminating the measurement of the DNA concentration.

ACKNOWLEDGMENTS

This study was funded by China National Funds for Distinguished Young Scientists (31125020), and

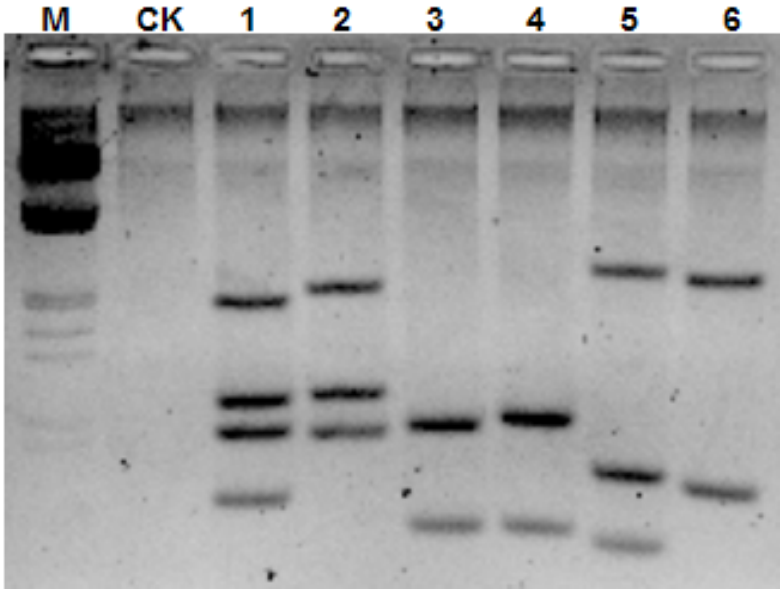

Figure 7. Southern blotting identifications of GM cotton plants. M, molecular marker; CK, DNA from non-transgenic cotton plant; 1 to 4, DNA from pollen tube pathway; 5 to 6, from gene gun bombardment mediated transgenic plants, respectively.

Innovation Scientists and Technicians Troop Construction Projects of Henan Province. We thank Prof. Chee-Kok Chin (Rutgers University, New Jersey, US) and Prof. Yuxian Zhu (Beijing University, Beijing,China) for advices on the manuscript.

Abbreviations

qRT-PCR, Quantitative real-time polymerase chain reaction; **Sad I,** stearoyl-acyl carrier protein desaturase I.

REFERENCES

Armour JA, Sismani C, Patsalis PC, Cross G (2000). Measurement of locus copy number by hybridization with amplifiable probe. Nucleic Acids Res. 28 (2): 605-609.

Kim SR, Lee J, Jun SH, Park S, Kang HG, Kwon S, An G (2003). Transgene structures in T-DNA-inserted rice plants. Plant Mol. Biol. 52: 761-773.

Lucito R, West J, Reiner A, Alexander J, Esposito D, Mishra B, Powers S, Norton L, Wigler M (2000). Detecting gene copy number fluctuation in tumor cells by microarray analysis of genomic representations. Genome Res. 10 (11): 1726–1736.

Makarevitch I, Svitashev SK, Somers DA (2003). Complete sequence analysis of transgene loci from plants transformed via microprojectile bombardment. Plant Mol. Biol. 52: 421-432.

Mason G, Provero P, Vaira AM, Accotto GP (2002). Estimating the number of integrations in transformed plants by quantitative real-time PCR. BMC Biotechnol. 2: p. 20.

Pfaffl MW (2001). A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 29: e45.

Paterson AH (1993). A rapid method for extraction of cotton Genomic DNA suitable for RFLP or PCR analysis. Plant Mol. Biol. Rep. 1: 122-127.

Song P, Cai CQ, Skokut M (2002). Quantitative real-time PCR as a screening tool for estimating transgene copy number in WHISKERS (TM)-derived transgenic maize. Plant Cell Rep. 20: 948-954.

Weng HB, Pan AH, Yang LT (2004). Estimating number of transgene copies in transgenic rapeseed by real-time PCR assay With HMG I/Y as an endogenous reference gene. Plant Mol. Biol. Rep. 22: 289-300.

Yang L, Ding J, Zhang C, Jia J, Weng H, Liu W, Zhang D (2005a). Estimating the copy number of transgenes in transformed rice by real-time quantitative PCR. Plant Cell Rep. 23: 759-763.

Yang L, Pan A, Zhang K, Yin C, Qian B, Chen J, Huang C, Zhang D (2005b). Qualitative and quantitative PCR methods for event-specific detection of genetically modified cotton Mon1445 and Mon531. Transgenic Res. 14: 817-831.

Zhang J, Cai L, Cheng J, Mao H, Fan X, Meng Z, Chan KM, Zhang H, Qi J, Ji L, Hong Y (2008). Transgene integration and organization in cotton (*Gossypium hirsutum L.*) genome. Transgenic Res. 17: 293-306.